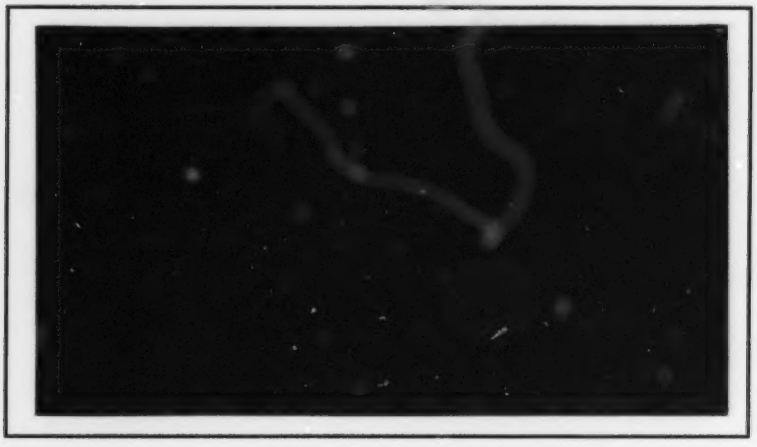

CENTRE FOR
SYSTEMS
SCIENCE

LCCR

Simon Fraser University, Burnaby, B.C., Canada V5A 1S6 / (604) 291-4704

Adaptive Compensation For Imbalance and Offset Losses In Direct Conversion Transceivers

James K Cavers and Maria Liao

CSS/LCCR TR91-15

December 1991

ADAPTIVE COMPENSATION FOR IMBALANCE AND OFFSET LOSSES IN DIRECT CONVERSION TRANSCEIVERS

James K. Cavers
Communications Science Lab
School of Engineering Science
Simon Fraser University
Burnaby B.C. V5A 1S6
Canada

and

Maria Liao
Glenayre Electronics Ltd
1570 Kootenay Street
Vancouver B.C. V5K 5B8
Canada

ABSTRACT

The current interest in linear modulation and multilevel signals has resulted in an emphasis on DSP implementations to achieve precision signal manipulation. However, most transceivers, and direct conversion designs in particular, rely on analog implementations of the quadrature modulator and demodulator, thereby sacrificing much of the precision gained through DSP. This paper focuses on the three principal impairments of analog quadrature modulators and demodulators: gain imbalance, phase imbalance, and DC offset.

The paper contains three main contributions. First is an analysis and quantitative assessment of the losses - primarily a degraded BER and out-of-band power in the transmitted signal - due to imbalances and offsets. The second contribution is an adaptive compensation technique for the quadrature modulator at the transmitter, and the third is a compensation technique for the quadrature demodulator at the receiver. Both compensation methods converge quickly, and present only a modest computational load.

Abbreviated version first presented at IEEE Vehicular Technology Conference, St. Louis, 1991.

This work was supported in part by the British Columbia Advanced Systems Institute.

December, 1991

I INTRODUCTION

There has been much recent effort in new designs for transceivers used in mobile communications. The general approach is to combine RF functions with DSP to allow linear modulation techniques and permit flexibility of modulation format and receiver processing. A favoured configuration at the the transmitter is baseband generation of the I and Q components by DSP, followed by quadrature modulation to IF or RF. At the receiver, a quadrature demodulator recovers the I and Q components from an RF or IF signal, so they can be processed further by DSP. Clearly, the quadrature modulator and quadrature demodulator (quad mod and quad demod, for brevity) are at the center of the new transceiver designs.

Analog implementations of a quad mod and quad demod suffer from several deficiencies [1,2], of which amplitude and phase imbalance and DC offset are the most important. Although all-DSP quad mod and demod designs [3,4] eliminate these problems, they do so at the expense of power consumption and a relatively low frequency IF interface. This paper focusses on an alternative design approach which retains the analog quad mod and demod, with their wide bandwidth and low power operation, and compensates for their impairments by DSP algorithms.

The paper contains three main contributions. First is an analysis and quantitative assessment of the impairments due to analog implementations. These problems include spurious tones and intermod products at the transmitter and a degraded BER due to distortion of the signal constellation. The asymptotic degradation in BER performance in 16QAM, for example, is shown to be 0.82 dB for a 5° phase imbalance, and 0.65 dB for 5% gain imbalance, and these imbalance figures are difficult to achieve in the transmitter and receiver combination. The second contribution is an adaptive procedure at the transmitter for compensation of the quad mod imbalance and offset losses. It features lower computational load and faster convergence than a previously reported compensator [5]. The third contribution is an adaptive compensation method at the receiver for minimization of losses due to both quad mod and demod. The paper demonstrates that the LMS algorithm can be used to adapt the receiver compensator, in a manner similar to an equalizer, and that the

computational load is equivalent to that of a short (3 tap) equalizer.

II SYSTEM MODEL

A. *Ideal Model*

Figure 1 presents a simplified model of the direct conversion transmitter and receiver. At the transmitter, the cascade of transmit modem (TM), predistorter (PD), quad mod compensator (QMC) and quad mod (QM) produces a real bandpass signal $\tilde{v}_q(t)$ as input to the power amplifier (PA). The PD is intended to linearize the PA, and the QMC corrects for imbalance and offset errors in the quad mod. Adaptation of the compensator requires only an envelope detector (ED) in the feedback path. At the receiver, the front end provides low noise amplification (LNA) and enough bandpass filtering (BPF) that the quad demod (QDM) does not operate in significantly nonlinear regions. A matched filter (MF) also provides adjacent channel rejection, and the quad demod compensator (QDMC) corrects for the QDM distortions prior to detection in the receive modem (RM).

Figure 2 shows the ideal model of a quad mod and quad demod. For the moment, we ignore the effects of the predistorter and the quad mod compensator, so that the desired modulation is represented by the complex envelope $v_m(t)$. The I and Q (real and imaginary) components of $v_m(t)$ are applied to double balanced mixers fed by cosine and sine references, to produce the bandpass signal $\tilde{v}_q$ with complex envelope $v_q(t)$. Generally, we will use a tilde to indicate a real bandpass signal corresponding to a complex envelope. Ideally, $v_q(t) = v_m(t)$, so that

$$\begin{aligned}\tilde{v}_q(t) &= \mathrm{Re}\left[v_q(t)\ \exp(j2\pi f_c t)\right] \\ &= \mathrm{Re}\left[v_m(t)\ \exp(j2\pi f_c t)\right] \end{aligned} \tag{1}$$

The receiver recovers the complex envelope of the received signal $\tilde{v}_r(t)$ by its own mixers and cosine and sine references. The lowpass filters remove the double frequency components, and may be used to reject much of the adjacent channel interference. Its output $v_d(t) = v_r(t)$, ideally.

Between transmitter and receiver, the signal passes through a bandpass channel with noise and distortion. For the purposes of this paper, it is sufficient to include only white noise and a phase rotation $\theta(t)$, that represents the difference in phase between transmitter and receiver

oscillators, so that

$$v_r(t) = v_q(t)e^{j2\pi\theta(t)} + n(t) \tag{2}$$

In the absence of noise, phase shift and imbalance and offset distortions, we have $v_d(t) = v_m(t)$, as desired.

B. Quadrature Modulator

A more realistic model of an analog quad mod is shown in Figure 3. The I and Q components entering from the DSP-based quad mod compensator pass through a DAC and active filter combination, which results in DC offset and gain differences between the two branches. The phase splitter at the local oscillator does not normally produce an exact $\pi/2$ separation, so the cosine and sine references are inexact. The mixers are never perfectly balanced, again causing gain differences. In addition, some of the LO signal can leak into the output signal $\tilde{v}_q(t)$, with an effect similar to DC offset.

We can summarize the quad mod errors as follows: DC offset errors of a_1 and a_2, amplitude gains in the I and Q branches of α and β, respectively, and a phase splitter error (i.e. the difference from $\pi/2$) of ϕ. We can further define the gain ratio γ and gain imbalance ε as

$$\gamma = \alpha/\beta \qquad \text{and} \qquad \varepsilon = \gamma - 1 \tag{3}$$

Typical values achievable with careful design are a gain imbalance of 2-3%, a phase imbalance of 2-3°, and a DC offset of about 2-3% of full scale.

To write the analog quad mod transformation concisely, we represent a complex envelope as a length-2 column vector containing the real and imaginary components. For example, $v_m(t))$ is represented as $\mathbf{v}_m(t) = \left(v_{m1}(t), v_{m2}(t)\right)^T$. The quad mod effect is then

$$\mathbf{v}_q(t) = \mathbf{M}\left(\mathbf{v}_m(t) + \mathbf{a}\right) \tag{4}$$

where the matrix $\mathbf{M}$ and vector $\mathbf{a}$ are defined as

$$\mathbf{M} = \begin{bmatrix} \alpha\cos(\phi/2) & \beta\sin(\phi/2) \\ \alpha\sin(\phi/2) & \beta\cos(\phi/2) \end{bmatrix} \qquad \mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \tag{5}$$

For symmetry, the phase splitter error has been distributed equally between

I and Q channels. An asymmetric model is considered in Section V.

We also adopt a power normalization, in order not to confuse effects due to distortion in the quad mod with ones due to an overall change in power. If the desired transmitted signal is a CW tone f_o Hz away from channel center:

$$v_m(t) = \exp(j2\pi f_o t) \tag{6}$$

or if the I and Q components are statistically independent, equal variance signals, then it is easily shown that the condition $\alpha^2 + \beta^2 = 2$ leaves the power of the information-bearing part of the signal unchanged. In terms of γ, therefore, we have

$$\alpha = \gamma\sqrt{\frac{2}{1+\gamma^2}} \qquad \beta = \sqrt{\frac{2}{1+\gamma^2}} \tag{7}$$

C. Quadrature Demodulator

Figure 3 also shows a physical model of the quad demod. As in the modulator, the phase split between cosine and sine references differs from the ideal $\pi/2$, and variations between mixers and between active filters result in amplitude gain differences in the I and Q branches. In addition, the baseband I and Q outputs contain a DC offset, incurred largely in the active filters, but with contributions from self-mixing of the LO output in the mixers, and from 1/f noise [2]. Thus the quadrature demodulator can be represented by

$$\mathbf{v}_d(t) = \mathbf{D}\,\mathbf{v}_r(t) + \mathbf{b} \tag{8}$$

where **D** and **b** are similar to **M** and **a**, respectively, but with different parameter values.

D. Overall Model

The overall system - quad mod, channel and quad demod - is shown in Figure 4. The phase shift in the channel can be represented by the rotation matrix

$$\mathbf{R}(t) = \begin{bmatrix} \cos(\theta(t)) & -\sin(\theta(t)) \\ \sin(\theta(t)) & \cos(\theta(t)) \end{bmatrix} \tag{9}$$

Receivers often implement carrier phase tracking by direct adjustment of LO phase. In this case, $\mathbf{R}(t)$ is constant, and equal to the identity matrix $\mathbf{I}$ if the phase error is zero.

The overall effect on the signal and channel noise can be summarized by the following:

$$\mathbf{v}_d(t) = \mathbf{H}(t)\mathbf{v}_m(t) + \mathbf{c}(t) + \mathbf{m}(t) \tag{10}$$

where

$$\mathbf{H}(t) = \mathbf{D}\mathbf{R}(t)\mathbf{M} \qquad \mathbf{c}(t) = \mathbf{H}(t)\mathbf{a} + \mathbf{b} \tag{11}$$

and

$$\mathbf{m}(t) = \mathbf{D}\mathbf{n}(t) \tag{12}$$

The covariance matrix of $\mathbf{n}(t)$ is $N_o\mathbf{I}$ (where $\mathbf{I}$ is the identity matrix) if its power spectrum is flat and symmetric about f_c; however, the covariance matrix of $\mathbf{m}(t)$ is

$$\mathbf{C}_m = N_o\mathbf{D}\mathbf{D}^T \tag{13}$$

Note that power normalization at the receiver preserves the trace of $\mathbf{C}_m$, that is, the sum of real and imaginary variances.

The distortions of the quad demod may precisely compensate for that of the quad mod. A more conservative view is that they combine unhelpfully. The approximate overall distortion with careful analog design is therefore a gain imbalance of 5%, a phase imbalance of 5°, and a DC offset of 5% of full scale.

III EFFECTS ON TRANSMITTED SIGNAL

A. *Spurious Tones*

One effect of an analog quad mod is seen when the desired signal is a CW tone f_o Hz away from channel center, as in (6). The actual output signal consists of the desired frequency with slightly different power, plus a spurious tone due to the imbalance at $-f_o$ from the center, and another due to DC offset at the center. It is readily shown that the power ratio of the spurious tone at $-f_o$ to the desired tone at f_o is given by

$$\frac{P_{spur}}{P_{des}} = \frac{\gamma^2 + 1 - 2\gamma\cos\phi}{\gamma^2 + 1 + 2\gamma\cos\phi} \approx (\varepsilon^2 + \phi^2)/4 \qquad (14)$$

where the approximation holds for small imbalances. For combined gain and phase imbalances of 3% and 3°, respectively, the spurious tone is only 30.4 dB below the desired one.

In many situtations, spurious tones at this level are not a problem, since they are in the assigned channel. However, if the quad mod is used for SSB modulation in a configuration in which the Q channel is equal to the Hilbert transform of the I channel, then the tones fall into a neighbouring channel. Fortunately, the Weaver method of SSB generation eliminates this more serious problem.

B. *Intermodulation Products*

Less familiar problems arise when the analog quad mod output feeds a nonlinear power amplifier. The quad mod distortions can interact with the amplifier nonlinearity to produce unexpected intermodulation (IM) products.

One such situation is the single CW tone (6), as a simple representative of constant envelope modulations like FM. A constant envelope modulation is often used because it produces no IM products, even when the nonlinear amplifier is near saturation. However, the complex envelope $v_q(t)$ of the quad mod output follows an offset elliptical trajectory, rather than a centered circular one. The result is envelope fluctuation at the rate $2f_o$, and therefore additional amplitude and phase

distortion at the amplifier output. Figure 5 shows that we can generate a significant level of in-band IM products from a single tone. The Class AB amplifier is the one described in [6], operated at an output backoff of about 0.2 dB, and the quad mod imbalances and offsets are equal to 2%, 2° and 2%, respectively.

In another example, we include the linearizing predistorter preceding the quad mod. Although a mapping type predistorter [7] can compensate for the quad mod deficiencies, the more economical gain-based predistorter [6] cannot, resulting in residual intermodulation products. We simulated a perfectly adjusted gain-based predistorter working through a quadrature modulator with 2% gain imbalance, 2° phase imbalance and 2% DC offset, into the same amplifier as above. The output backoff (saturated power less peak output power) is about 0.22 dB for the 16QAM signal. As shown on Figure 6, the spectrum acquires IM skirts as much as 10 dB above the original spectral floor.

The damage is more severe in the case of an adaptive predistorter that employs quadrature demodulation of the amplifier output in the feedback path and adaptation by the secant method. Figure 6 shows that allowing both the quad mod and the quad demod to have imbalances and offsets of 2%, 2° and 2%, respectively, causes the IM floor to rise by about 25 dB. The design effort and cost of the adaptive predistorter is largely wasted as a result of the quad mod impairments.

IV EFFECTS ON SYMBOL ERROR RATE

Since the analog quad mod and quad demod distort the signal constellation, we can expect a degraded symbol error rate (SER). In this section, we quantify the loss for QAM systems.

A. Detection Model

We assume a matched filter following the quad demod and ideal sampling, so that symbol-by-symbol decisions are made on a received two-dimensional vector. The nominal signal points are defined on odd integers:

$$s = \left(\pm(2i+1),\ \pm(2k+1)\right), \quad i,k = 0,\ 1,\ldots. \tag{15}$$

The decision boundaries are therefore at even integers, vertically and horizontally, defining square cells of width 2 about each nominal signal point. The energy per symbol, defined as

$$E_s = \frac{1}{2}\, E[|s|^2] \tag{16}$$

takes the values 1 for QPSK and 5 for 16QAM.

B. Direct Calculation

Two effects change the SER from that obtained with perfect quad mod and demod. First is distortion of the constellation: if s_i is the i^{th} signal point, then its position error is

$$e_i = (H - I)s_i + c \tag{17}$$

Clearly, the position error is worse at larger constellation points. The second effect is the change in real and imaginary noise variances introduced by the quad demod, though their sum is invariant, as noted after (13). Applying the union bound to a typical signal point in a four-walled cell, we have

$$SER_i \le \frac{1}{4}\Big[Q\big((1-e_{i1})/\sigma_{m1}\big) + Q\big((1+e_{i1})/\sigma_{m1}\big) + Q\big((1-e_{i2})/\sigma_{m2}\big) + Q\big((1+e_{i2})/\sigma_{m2}\big)\Big] \tag{18}$$

with the obvious modification for cells with fewer walls. Figure 7 shows how the SER bound, averaged over all signal points for 16QAM, depends upon the imbalance and offset parameters. For simplicity of presentation, all the impairments have been ascribed to the quad demod (i.e. $\mathbf{M} = \mathbf{I}$, $\mathbf{a} = \mathbf{0}$). At a BER of 10^{-5}, combined impairments of 3%, 3° and 3% result in a 1 dB loss, and the loss increases at lower error rates.

C. Asymptotic Loss

From (18), for every error term that increases, there is another that decreases. However, as the SNR becomes large the asymptotic degradation is determined by the increased term; in fact, ultimately it is determined by the worst such increase in the constellation. In the simplified case where only the quad mod has imbalance impairments, we have $\sigma_{m1} = \sigma_{m2}$ and an asymptotic loss of effective SNR given by

$$L_{dB} = \max_{i,k} 20 \log(1-|e_{ik}|) \approx 8.7 \max_{i,k} |e_{ik}| \qquad (19)$$

though care must be taken that the max takes place at a true decision boundary, since some cells have fewer than four sides. From (17), the maximum loss is generally associated with the outer ring of signal points.

For gain imbalance only and 16QAM, substitution of (5) into (17) and (19) and series expansion shows that the loss is $L_{dB} \approx 13.1\varepsilon$, or 0.66 dB for $\varepsilon = 5\%$. Similarly, for phase imbalance only, the loss is $L_{dB} \approx 13.1\phi$, or 1.1 dB for 5° imbalance. Finally, for offset only, the loss is $L_{dB} = 8.7 \max_{k}(c_k)$, or 1.3 dB for offset equal to 5% of full scale.

V ADAPTIVE COMPENSATION AT THE TRANSMITTER

The two previous sections demonstrated that imperfections in the quad mod lead to significant levels of intermodulation products and a degraded BER. Here we demonstrate that adaptive DSP-based correction of the impairments is computationally economical and converges rapidly.

A. Asymmetric Form

As in an earlier work [5], we employ only an envelope detector in the feedback path. However, our approach differs in that we use a more economical asymmetric form of compensator, and we demonstrate rapid convergence of the adaptation, so that less time is spent on test signals. In fact, only 20 to 40 measurements are required for complete adaptation.

It is clear from (4) that the compensator has the form

$$\mathbf{v}_c(t) = C\mathbf{v}_m(t) + \mathbf{f} \tag{20}$$

where, ideally, $C = M^{-1}$ and $\mathbf{f} = -\mathbf{a}$, and we have ignored the predistorter for simplicity. Substituting (5) leads to the symmetric compensator of [5]. On the other hand, we could equally well have attributed the entire phase imbalance error to the sine branch. In this case

$$M = \begin{bmatrix} \alpha & \beta \sin(\phi) \\ 0 & \beta \cos(\phi) \end{bmatrix} \quad \text{and} \quad M^{-1} = \alpha^{-1} \begin{bmatrix} 1 & -\tan(\phi) \\ 0 & \gamma \sec(\phi) \end{bmatrix} \tag{21}$$

One measure of complexity is the number of instructions required, using a common DSP CPU, the TMS320C2X. Executing (20) with the symmetric form uses 16 instructions, compared with 13 in the asymmetric form (21).

Further simplification is possible if the division by α in (21) is neglected on the grounds that it just produces an overall scale change of a few per cent. The cost is then only 11 instructions. We will adopt this simplified compensator, as illustrated in Figure 8. Adaptation consists of adjusting the four remaining coefficients c_{12}, c_{22}, f_1 and f_2 in (20) until there is no imbalance or offset at the quad mod output.

B. Adaptation of Offset

Eliminating DC offset is accomplished by setting $\mathbf{v}_m(t)=0$ and adjusting

$\mathbf{f}$ to minimize the envelope detector output v_e. Because of the diode characteristics in an envelope detector, v_e describes a locally quadratic surface near the minimum

$$v_e(\mathbf{f}) = G_e \|\mathbf{v}_q\|^2 + B_e \simeq G_e \|\mathbf{f}+\mathbf{a}\|^2 + B_e \tag{22}$$

with four unknowns: the minimum value B_e, which depends on bias in the envelope detector; the gain around the loop G_e, which is known only nominally; and the two components of $\mathbf{a}$. Note that $\mathbf{M}$ has been approximated as the identity matrix, since gain and phase imbalance have only a second order effect in the estimation of the minimum point. For simplicity, we also write

$$v_e(\mathbf{f}) = \mathbf{h}(\mathbf{f})^T \mathbf{p} \tag{23}$$

where

$$\mathbf{h}(\mathbf{f}) = \left[f_1^2 + f_2^2,\ f_1,\ f_2,\ 1\right]^T$$

$$\mathbf{p} = \left[G_e,\ 2G_e a_1,\ 2G_e a_2,\ G_e(a_1^2 + a_2^2) + B_e\right]^T \tag{24}$$

Estimation of $\mathbf{a}$, and hence the optimum $\mathbf{f}$, is a two-step process. First, make N measurements $m(k)$, $k=1,..,N$ of v_e at different values $\mathbf{f}(k)$, and estimate $\mathbf{p}$ from them. A suitable approach is to minimize the sum of squared errors between the measurements and the fitted surface. The optimum estimate of $\mathbf{p}$ is easily shown to be

$$\hat{\mathbf{p}} = \mathbf{R}^{-1}\mathbf{w} \tag{25}$$

where

$$\mathbf{R} = \frac{1}{N}\sum_{k=1}^{N} \mathbf{f}(k)\ \mathbf{f}^T(k) \qquad \mathbf{w} = \frac{1}{N}\sum_{k=1}^{N} m(k)\ \mathbf{f}(k) \tag{26}$$

In the second step, we obtain the estimated optimum $\hat{\mathbf{f}}$ from $\hat{\mathbf{p}}$ using (24):

$$\hat{f}_1 = -\frac{1}{2}\ \frac{\hat{p}_2}{\hat{p}_1} \qquad \hat{f}_2 = -\frac{1}{2}\ \frac{\hat{p}_3}{\hat{p}_1} \tag{27}$$

In principle, with only the four unknowns B_e, G_e, a_1 and a_2 to determine, only four measurements are required. In practice, given measurement noise and small departures from the ideal quadratic surface, a few more are required. An effective method is to make four initial

measurements with $\mathbf{f}(k)$ selected as the corners of a square centered on the origin. Subsequent selections for $\mathbf{f}(k)$ are the estimated optimum values $\hat{\mathbf{f}}(k-1)$ derived from all previous measurements. No more than about eight measurements should be required at reasonable levels of measurement noise. This is considerably faster than the method in [5], which requires alternating minimizations with respect to f_1 and f_2.

C. Adaptation of Gain and Phase Imbalance

Adaptation of gain and phase imbalances takes place after the offset adaptation. Previous work [5] showed that the symmetric compensator allows independent adjustment of gain, then phase. In contrast, the present work uses the asymmetric form, in which gain and phase are coupled. An algorithm for simultaneous adaptation is given below.

From (21), the compensator matrix can be written

$$\mathbf{C} = \begin{bmatrix} 1 & -\tan(\phi)+e_p \\ 0 & \gamma \sec(\phi)+e_g \end{bmatrix} \tag{28}$$

where e_p and e_g are the misadjustment errors. The resulting composite compensator/quad mod matrix is then

$$\mathbf{M}\,\mathbf{C} = \beta \begin{bmatrix} 1+\varepsilon & (1+\varepsilon)e_p+e_g\sin(\phi) \\ 0 & 1+\varepsilon+e_g\cos(\phi) \end{bmatrix} \tag{29}$$

Following [5], as test signals we use tones at channel center with four different phases. With amplitude A, they are

$$\mathbf{v}_m^{<1>} = [0,A]^T \qquad \mathbf{v}_m^{<2>} = [A,0]^T$$

$$\mathbf{v}_m^{<3>} = [A,A]^T/\sqrt{2} \qquad \mathbf{v}_m^{<4>} = [-A,A]^T/\sqrt{2} \tag{30}$$

Adaptation is complete when all four phases give the same output v_e from the envelope detector. Error signals to drive the adaptation are obtained by subtracting test signal outputs in pairs

$$f_g(e_g,e_p) = G'_e \|\mathbf{v}_q^{<1>}\| - G'_e \|\mathbf{v}_q^{<2>}\|$$

$$f_p(e_g,e_p) = G'_e \|v_q^{<3>}\| - G'_e \|v_q^{<4>}\| \tag{31}$$

where A is assumed to be large enough that the envelope detector is now operating in a more linear region than for offset adjustment, and G'_e is the the gain around the loop.

Substitution of (29) and (30) into (31) and series expansion to first order shows that

$$f_g(e_g,e_p) \approx G'_e A\, e_g \qquad f_p(e_g,e_p) \approx G'_e A\, e_p \tag{32}$$

even for the asymmetric compensator. An appropriate iteration algorithm for adjustment of the second column of C is therefore

$$c_{12}(k+1) = c_{12}(k) - \delta\, f_p(e_g(k),e_p(k))$$

$$c_{22}(k+1) = c_{22}(k) - \delta\, f_g(e_g(k),e_p(k)) \tag{33}$$

where δ is a step size parameter. Selection of δ is straightforward. From (32), it must satisfy

$$0 < \delta\, G'_e\, A < 2 \tag{34}$$

for convergence. Since the difference (31) is sensitive to measurement noise, $\delta\, G'_e\, A$ should be selected to be small and positive, to provide an averaging effect. In this case, the time constant is about $(\delta\, G'_e\, A)^{-1}$ iterations. In practice, there is little noise in this type of power measurement, so that we need fewer than about 8 iterations to reduce the imbalances by an order of magnitude. Since each iteration requires 4 power measurements, we need only about 32 power measurements for satisfactory performance. If there were no noise, we would still need about 3 iterations (12 power measurements), because G'_e is typically known only to about 10%. Thus only 12 to 32 measurements are necessary for adaptation of imbalances.

The same considerations would apply to adaptation of the symmetric form of compensator. Since the gain and phase error measurements are uncoupled, at least to first order (32), the convergence speed of the asymmetric compensator presented here is no less than that of the symmetric compensator.

VI ADAPTIVE COMPENSATION AT THE RECEIVER

If the distortion of the constellation were corrected, then the loss of effective SNR would be reduced or eliminated. This section outlines an adaptive technique by which the receiver can correct distortion which is initially unknown or changes during transmission. The compensator resembles an adaptive equalizer, although the correction itself is memoryless, unlike an FIR filter.

A. Adaptation Iteration

As suggested by Figure 9, a transformation like those in the quad mod and demod can eliminate the distortion. Decisions are based on the corrected variable $\mathbf{u}(k) = \mathbf{Q}\mathbf{r}(k) + \mathbf{d}$, where from (10,15)

$$\mathbf{r}(k) = \mathbf{H}\mathbf{s}(k) + \mathbf{c} + \mathbf{D}\mathbf{n}(k) \tag{35}$$

It is more convenient for analysis to use the length-3 extended received vector $\mathbf{r}'$ and 3x2 matrix G:

$$\mathbf{r}'(k) = \left[\frac{\mathbf{r}(k)}{1}\right] \qquad \mathbf{G} = [\mathbf{Q}|\mathbf{d}]^T \tag{36}$$

so that the correction is written

$$\mathbf{u}(k) = \mathbf{G}^T\mathbf{r}'(k) \tag{37}$$

and the resulting error is

$$\mathbf{e}(k) = \mathbf{u}(k) - \mathbf{s}(k) \tag{38}$$

The fact that $\mathbf{r}(k)$ and 1 in (36) are not dimensionally consistent requires some care in analysis (see Sections *B* and *C* below).

Like an equalizer, the compensator can be operated in reference directed or decision directed mode. If decision directed, $\mathbf{s}(k)$ is replaced by the decision $\hat{\mathbf{s}}(k)$.

We want to minimize the error variance

$$\mathcal{E} = E[|\mathbf{e}(k)|^2] = E[e_1^2(k)] + E[e_2^2(k)] \tag{39}$$

with respect to G. Note that $e_1(k)$ and $e_2(k)$ depend only on the first and second columns $\mathbf{g}_1$ and $\mathbf{g}_2$, respectively, of G. They are therefore two independant minimization problems. The analysis below deals only with $e_1(k)$, since the other component is similar. The analysis is also brief,

since the LMS algorithm is well known from adaptive equalizer theory, e.g. [8].

First, define the second moment matrix of $\mathbf{r}'$:

$$C_{r'} = E[\mathbf{r}'(k)\mathbf{r}'^{T}(k)] = \left[\begin{array}{c|c} C_r & c \\ \hline c^T & 1 \end{array}\right] \tag{40}$$

which contains the 2x2 second moment matrix of $\mathbf{r}$

$$C_r = E_s \mathbf{H}\mathbf{H}^T + N_o \mathbf{D}\mathbf{D}^T + cc^T \tag{41}$$

Next, define the joint second moment of the received vector and the symbol being estimated:

$$w_1 = E[s_1(k)\mathbf{r}'(k)] = E_s \left[\frac{h_1}{0}\right] \tag{42}$$

where h_1 is the first column of $\mathbf{H}$.

The quantity to be minimized is

$$\begin{aligned} \mathcal{E}_1 &= E[e_1^2(k)] = E[(g_1^T \mathbf{r}'(k) - s_1(k))^2] \\ &= g_1^T C_{R'} g_1 - 2 w_1^T g_1 + E_s \end{aligned} \tag{43}$$

Since the gradient with respect to g_1 is

$$\nabla_{g_1} \mathcal{E}_1 = 2\, E[e_1(k)\mathbf{r}'(k)] = 2\,(C_{r'} g_1 - w_1) \tag{44}$$

the minimizing g_1 is given by

$$g_{1opt} = C_{r'}^{-1} w_1 \tag{45}$$

resulting in

$$\mathcal{E}_{1min} = E_s - g_{1opt}^T w_1 \tag{46}$$

The adaptation iteration adjusts g_1 with a step in the opposite direction to the estimated gradient:

$$g_1(k+1) = g_1(k) - \delta e_1(k)\mathbf{r}'(k) \tag{47}$$

where the step size parameter δ determines the tradeoff between speed of convergence and jitter in the converged state.

The computational load of this compensation technique is quite modest,

and is the same as that of a 3-tap complex linear equalizer.

B. Excess Mean Squared Error and Convergence Speed

Jitter about the converged value results in a value of $\mathcal{E}_1$ higher than $\mathcal{E}_{1min}$. It is shown in [8] that this "excess mean squared error" $\mathcal{E}_{1\delta}$ is approximated by

$$\mathcal{E}_{1\delta} = \delta\, \mathcal{E}_{1min}\, \mathrm{trace}(C_{r'})/2 \tag{48}$$

We saw in Section 4.2 that asymptotic loss is directly related to the error, so we might try to keep $\mathcal{E}_{1\delta}$ no more than 5% of $\mathcal{E}_{1min}$. We note also from (40), (41) that $C_{r'}$ is strongly diagonal, with a trace of approximately $2E_s+1$. Therefore

$$\delta \le 0.1/(2E_s+1) \tag{49}$$

Expected convergence speed is determined by the largest eigenvalue ρ_{max} of $I - \delta C_{r'}$. The time constant, in number of iterations, is

$$N_c = -1/\ln(\rho_{max}) \simeq 1/(1-\rho_{max}) \tag{50}$$

However, ρ_{max} is related to the minimum eigenvalue of $C_{r'}$ by $\rho_{max} = 1 - \delta\lambda_{min}$. From (40) and (41), with small distortion and reasonable SNR, the eigenvalues of $C_{r'}$ are approximately E_s, E_s and 1. Which one is the smallest depends on the arbitrary scale assigned in the signal processor to the dimensionless 1, compared to E_s. Here we will assume that $E_s \gg 1$, so that $\lambda_{min}=1$ and

$$N_c \simeq 1/\delta \ge 10(2E_s + 1) \tag{51}$$

where the inequality comes from (49). From (51), convergence time is roughly proportional to the ratio of eigenvalues, that is, E_s compared to unity.

The strongly diagonal nature of $C_{r'}$ suggests an improvement. Since the three components of g_1 are only weakly coupled, they adapt to a large extent independently, with the DC component slowest, since it is associated with λ_{min}. The step size parameter can therefore be made E_s times smaller for the first two components to improve their excess mean squared error contribution, with little change in overall convergence speed.

Equivalently, the AGC can be adjusted to put E_s at a value comparable to unity in the signal processor.

C. Numerical Results

One way to illustrate adaptation behaviour is with the convergence of mean squared error, obtained by substituting the sequence of iterations $g_1(k)$ into (43). Figure 10 illustrates typical decision directed convergence for 16QAM at an SNR giving a converged SER of about 1%. The average value of E_s is 10 times the value assigned to unity in the signal processor, and the step size parameter is the same for all three components. For simplicity of presentation, all the distortion has been attributed to the quad mod, and the quad demod is considered error-free. The initial values of gain and phase imbalance and offset are quite large, at 10%, 10° and 10%. The excess mean squared error is evident in the curve for $\delta = 0.01$. The adaptation is generally satisfactory, though its speed may be problematic in reception of short bursts if successive bursts may be from different sources.

VII CONCLUSIONS

Analog implementations of quadrature modulators and demodulators have deficiencies - primarily amplitude and phase imbalances and DC offset - that result in several troublesome problems for digital transceivers. At the transmitter, the quad mod can result in intermod products even for nominally constant envelope modulation. More generally, it can reduce or eliminate the benefits of linearizing predistorters. At the receiver, the combined effects of quad mod and quad demod cause an increase in error rate, or loss of noise margin. For 16QAM, the loss is about 1 dB even for analog impairments at a level requiring difficult design.

Our compensation technique at the transmitter is a more efficient asymmetric version of one presented earlier. We also provide fast new adaptation methods requiring a total of between 20 and 40 measurements for satisfactory convergence.

Our adaptive correction procedure at the receiver can reduce or eliminate the SER degradation. It requires only a modest increase in computational load for a DSP-based receiver, and can cope with large values of analog implementation error.

Since the compensation algorithms permit less precise design of the analog quad mod and demod, the overall cost of a transceiver may be lowered.

REFERENCES

[1] A. Bateman and D.M. Haines, "Direct Conversion Transceiver Design for Compact Low-Cost Portable Mobile Radio Terminals", Proc. IEEE Vehic. Techn. Conf., pp. 57-62, San Francisco, May 1989.

[2] P. Estabrook, and B.B. Lusignan, "The Design of a Mobile Radio Receiver using a Direct Conversion Architecture", Proc. IEEE Vehic. Techn. Conf., pp. 63-72, San Francisco, May 1989.

[3] J.K. Cavers and S.P. Stapleton, "A DSP-Based Alternative to Direct Conversion Receivers for Digital Mobile Communications", Proc. IEEE Global Telecommun. Conf., San Diego, December, 1990.

[4] S.P. Stapleton and J.K. Cavers, "Integrated DSP/RF Design for an MSAT Transmitter", Proc. Second Intl. Mobile Satellite Conf, Ottawa, June 1990.

[5] M. Faulkner, T. Mattson, W. Yates, "Automatic Adjustment of Quadrature Modulators", Electronics Letters, vol 27, no 3, pp 214-216, 1991.

[6] J.K. Cavers, "A Linearizing Predistorter With Fast Adaptation and Low Memory Requirements", IEEE Trans Vehicular Techn, vol 39, no 4, pp 374-382, November 1990.

[7] Y. Nagata, "Linear Amplification Technique for Digital Mobile Communications", Proc IEEE Vehic Tech Conf, pp 159-164, San Francisco, 1989.

[8] J.G. Proakis and J.H. Miller, "An Adaptive Receiver for Digital Signaling Through Channels With Intersymbol Interference", IEEE Trans Inform Th, vol IT-15, no 4, pp 484-497, July 1969.

ILLUSTRATIONS

Fig 1: Block Diagram of Transceiver

Fig. 2: Ideal Quadrature Modulator and Demodulator

Fig. 3: Physical Model of Analog System

Fig. 4: Matrix Model of Transmission System

Fig. 5: Analog Quad Mod Generation of Spurious Tones From CW Signal

Fig. 6: Analog Quad Mod Degradation of Perfect Predistortion

Fig. 7: Degradation of 16QAM Symbol Error Rate

Fig. 8: The Asymmetric Quad Mod Compensator

Fig. 9: Decision Directed Compensation at Receiver of Imbalance and Offset

Fig. 10: Convergence Behaviour of Adaptive Compensator for 16QAM

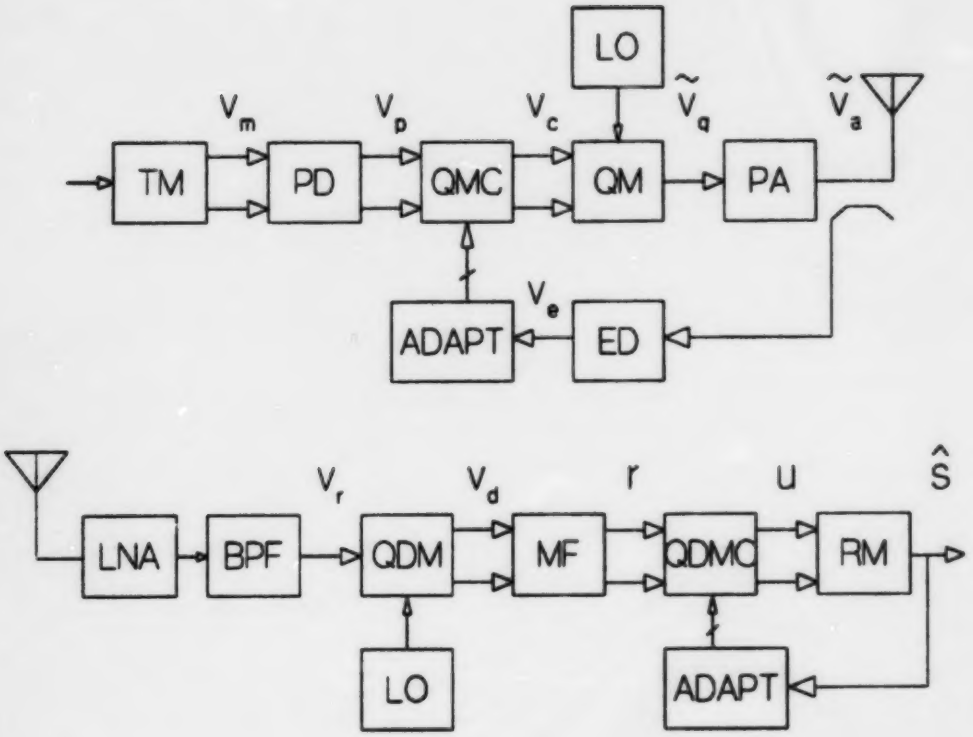

Figure 1: Block diagram of Transceiver

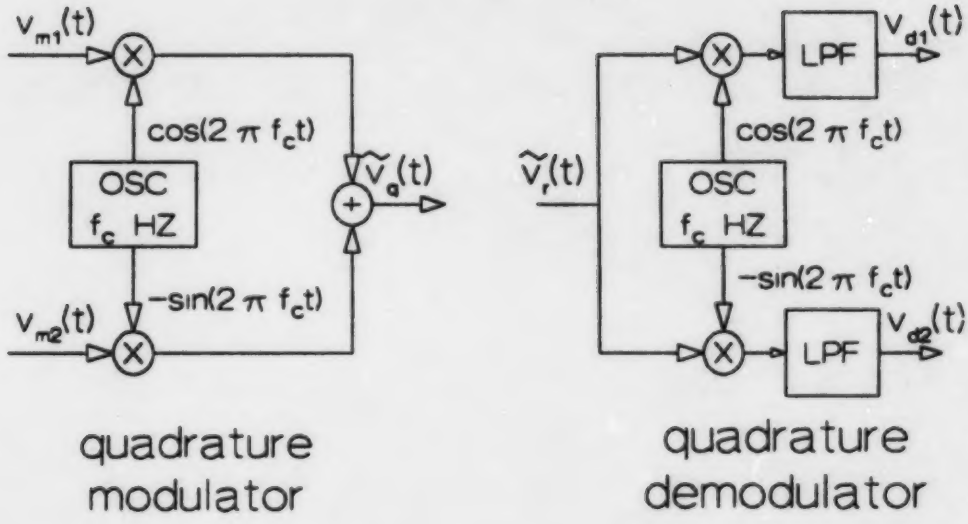

Figure 2: Ideal Quadrature Modulator and Demodulator

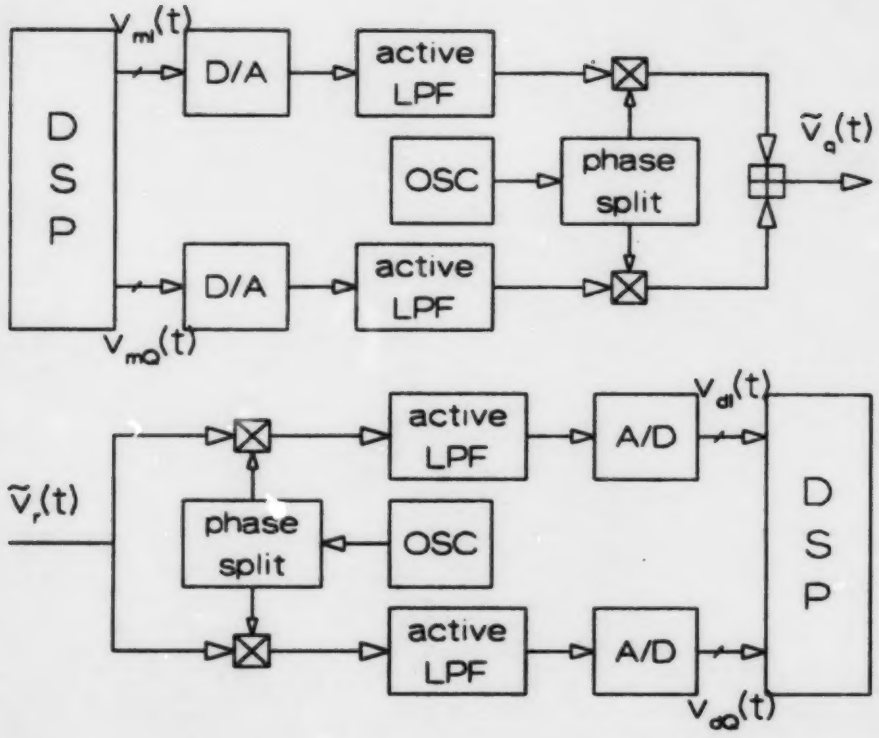

Figure 3: Physical Model of the Analog System

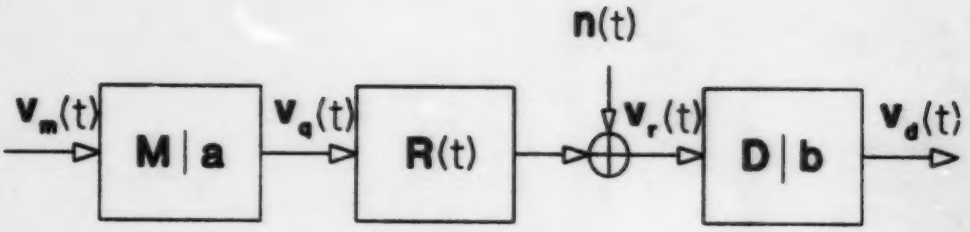

Figure 4: Matrix Model of the Transmission System

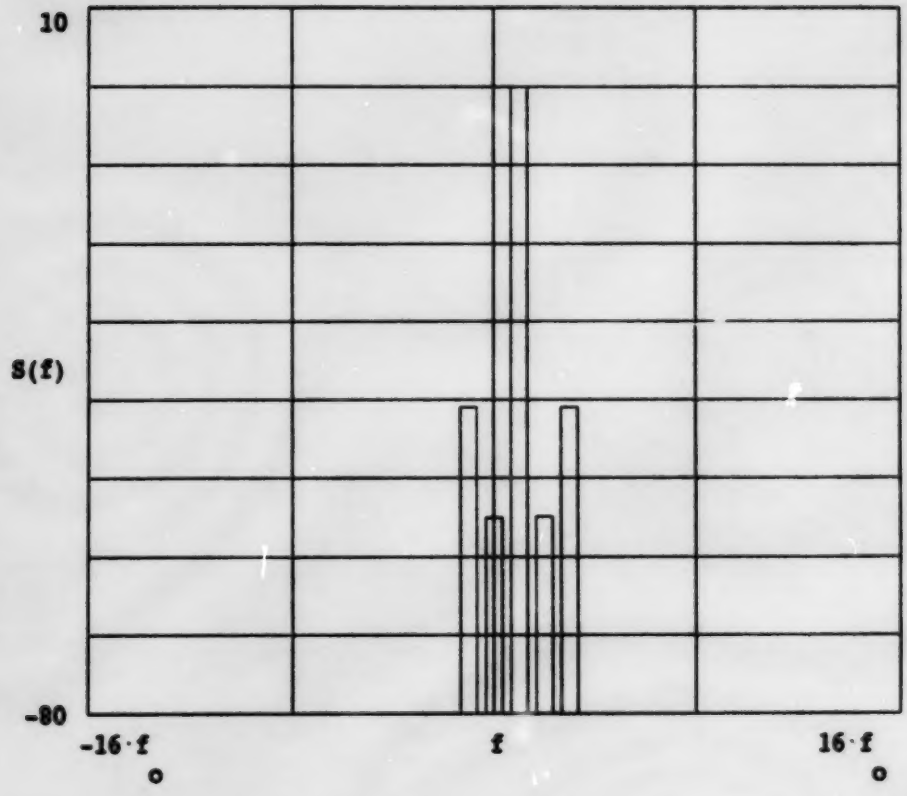

Figure 5: Spurious Tones Generated from CW by Analog Quad Mod and Nonlinear Amplifier

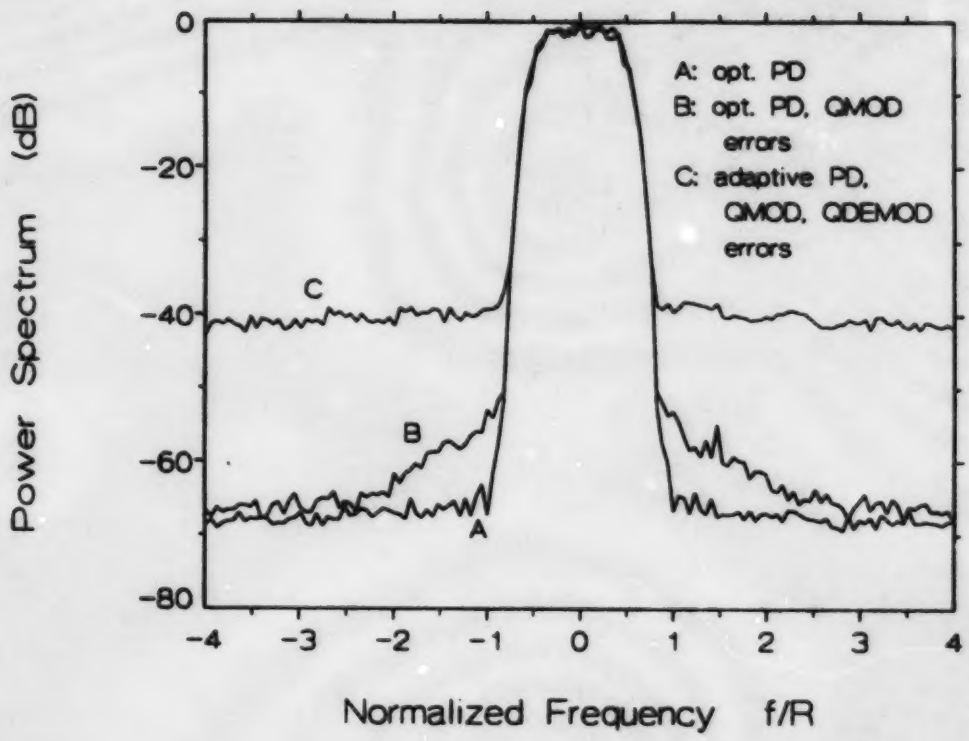

Figure 6: Analog Quad Mod Degradation of Perfect Predistortion

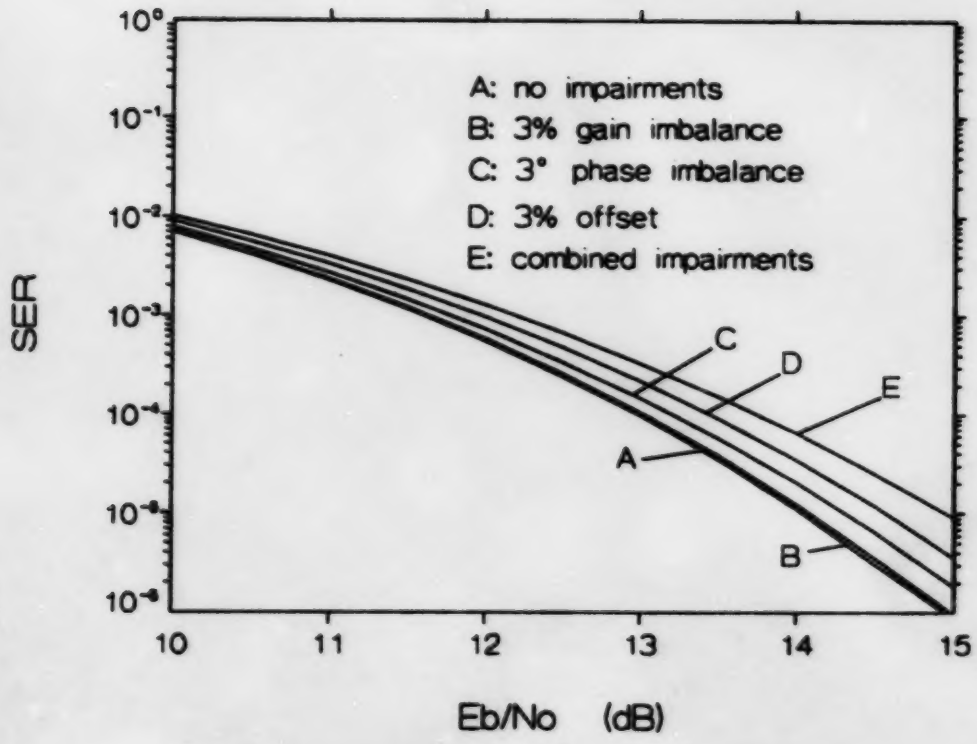

Figure 7: Degradation of 16QAM Symbol Error Rate

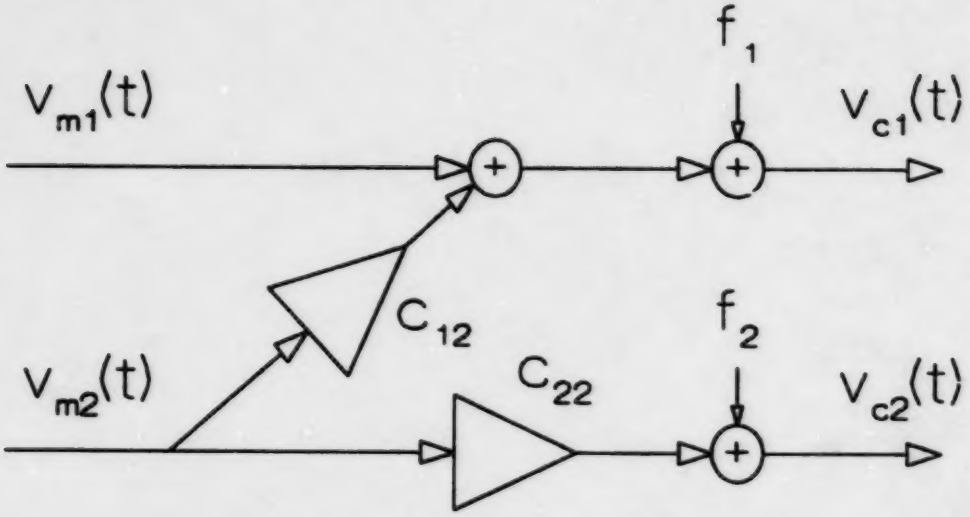

Figure 8: The Asymmetric Quad Mod Compensator

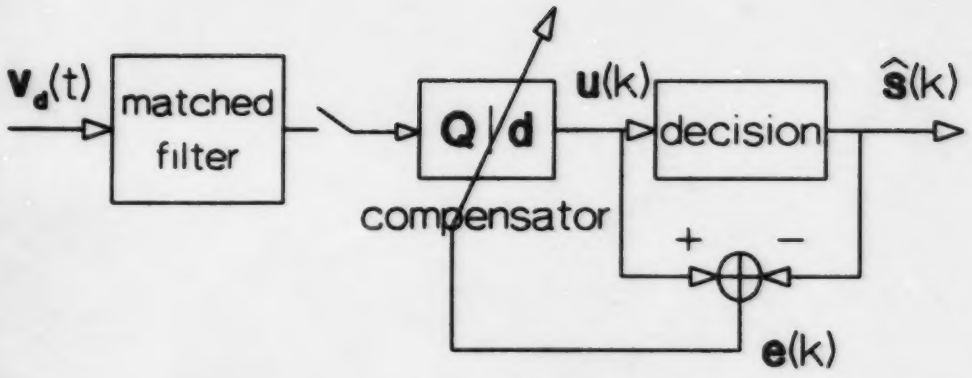

Figure 9: Decision Directed Quad Demod Compensation

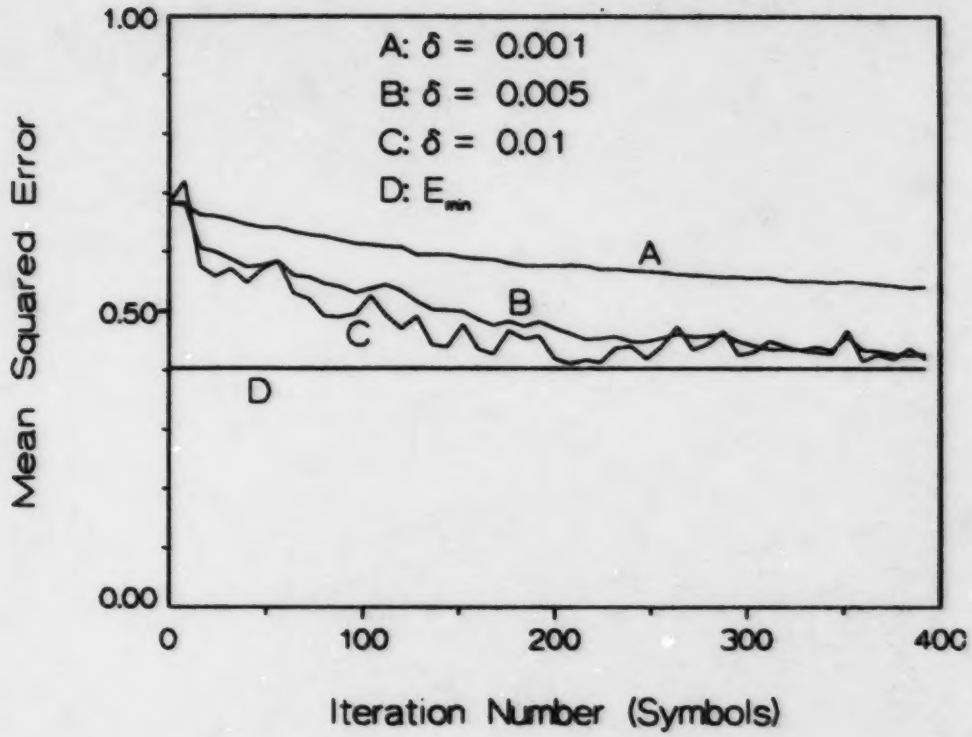

Figure 10: Convergence Behaviour of Quad Dumod Compensator for 16QAM

Centre for Systems Science

• Simon Fraser University established the **Centre for Systems Science** as a provincial centre for advanced study. The objective of CSS is twofold: coordinate and combine advanced systems research at SFU, and encourage application of research results within industry in order to further the technological development of British Columbia.

•CSS has assembled three research groups to streamline its efforts. The orientation of these groups is as follows: intelligent systems, microelectronics, and computer and communications systems. Because of shared research interests, this latter group merges with LCCR (**Laboratory for Computer and Communications Research**) for the publication of technical reports.

• CSS's three groups involve over 80 skilled faculty researchers principally from SFU's Schools of Engineering Science and Computing Science, but also draw substantially on Mathematics, Physics, Geography, Philosophy, Linguistics, Psychology, Communication, and Kinesiology.